

UNCLASSIFIED

AD NUMBER

AD826558

LIMITATION CHANGES

TO:

Approved for public release; distribution is unlimited.

FROM:

Distribution authorized to U.S. Gov't. agencies and their contractors; Critical Technology; SEP 1967. Other requests shall be referred to Air Force Technical Applications Center, Attn: VELA Seismological Center, Washington, DC 20333.

This document contains export-controlled technical data.

AUTHORITY

usaf ltr, 25 jan 1972

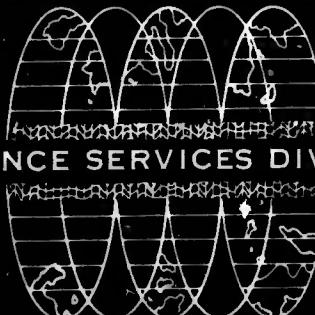
THIS PAGE IS UNCLASSIFIED

AD826558



**STATEMENT #2 UNCLASSIFIED**

This document is subject to special export controls and each  
transmittal to foreign governments or foreign nationals may be  
made only with prior approval of Chief, HEATAC, Wash, DC 20333



## SCIENCE SERVICES DIVISION



# TEXAS INSTRUMENTS INCORPORATED



VELA T/7701

This document is subject to special export controls and each  
transmittal to foreign governments or foreign nationals may  
be made only with prior approval of Chief, AFTAC  
Wash D.C. 20333

8 September 1967

**STATISTICS GOVERNING THE DESIGN AND  
PERFORMANCE OF NOISE-PREDICTION FILTERS**

**ADVANCED ARRAY RESEARCH**

**Special Report No. 3**

**Prepared by**

**John P. Burg      T.J. Cruise      Aaron H. Booker**

**George D. Hair, Program Manager**

**1-214-238-3473**

**TEXAS INSTRUMENTS INCORPORATED  
Science Services Division  
P.O. Box 5621  
Dallas, Texas 75222**

Contract No.	F33657-67-C-0708-P001
Contract Date	15 December 1966
Amount of Contract	\$625, 500
Contract Expiration Date	14 December 1967

**Sponsored by**

**ADVANCED ARRAY RESEARCH  
Nuclear Test Detection Office  
ARPA Order No. 624  
ARPA Program Code No. 7F10**



---

#### **ACKNOWLEDGMENT**

This research was supported by the  
**ADVANCED RESEARCH PROJECTS AGENCY**  
Nuclear Test Detection Office  
under Project VELA UNIFORM  
and accomplished under the technical direction of the  
**AIR FORCE TECHNICAL APPLICATIONS CENTER**  
under Contract No. F33657-67-C-0708-P001



---

## TABLE OF CONTENTS

Section	Title	Page
I	INTRODUCTION	I-1
II	MEAN-SQUARE ESTIMATION FORMULATION	II-1
III	INVARIANCE PROPERTIES OF $\alpha$ AND $\beta$	III-1
IV	PROBABILITY DENSITIES OF $\alpha$ AND $\beta$	IV-1
V	APPLICATIONS	V-1
	A. PROPERTIES OF $\alpha$	V-1
	1. Average Behavior	V-1
	2. Distribution Function	V-2
	3. Density Function	V-4
	B. CONFIDENCE-LEVEL PLOTS	V-4
	C. APPROXIMATIONS	V-4
	D. PROPERTIES OF $\beta$	V-10
	1. Average Behavior	V-10
	2. Density Function	V-12
	E. RESTRICTIONS AND EXTENSIONS	V-12
VI	GENERALIZED INVARIANCE PROPERTIES	VI-1
VII	REFERENCES	VII-1

## LIST OF APPENDIXES

Appendix	Title
A	DERIVATION OF PROBABILITY DENSITIES OF $\hat{\Gamma}$ , $\alpha$ , AND $\beta$
B	SUMMARY OF NOTATION



---

## LIST OF ILLUSTRATIONS

Figure	Description	Page
V-1	Cumulative Distribution of $\alpha$ for 12-Channel Array and 80 Samples	V-3
V-2	Probability Densities of $\alpha$ for $c = 12$ Channels	V-5
V-3	50-Percent Confidence Plot	V-6
V-4	80-Percent Confidence Plot	V-7
V-5	90-Percent Confidence Plot	V-8
V-6	99-Percent Confidence Plot	V-9
V-7	Probability Densities of $\beta$ for $c = 12$ Channels	V-13/14

## TABLE

Table	Title	Page
IV-1	Probability Densities of $\alpha$ and $\beta$	IV-2



---

## SECTION I

### INTRODUCTION

In certain applications of multichannel filtering techniques, the problem of signal extraction from a  $c$ -seismometer array reduces to one of noise prediction. For example, if the signal consists of vertical motion only and the array consists of a principal vertical seismometer and  $c-1$  horizontal seismometers, the horizontal seismometers serve only to predict the vertical noise on the principal seismometer. Thus, the horizontal seismometers minimize the vertical noise power on the principal seismometer without affecting the signal and thereby maximize the signal-to-noise ratio.

For the more general problem, the signal will be present on all seismometers; however, if the signal covariance matrix between channels is known, the  $c$ -channel outputs can be combined to form a new set of  $c$  outputs (channels) so that the signal will be present on only the number of channels equal to the rank of the signal covariance matrix. Once in this form, the channels without signals are used to reduce the noise power on the channels with signals. Typically, the channels with signals are assigned unity gain so that the signal extracted from the array outputs will be an unbiased estimate. Maximum likelihood processing under the zero-mean Gaussian noise assumption takes this form: the rank-1 signal is isolated on one channel and the remaining noise-only channels are employed to reduce the noise power on the signal channel. This report is restricted to rank-1 signals, although higher-rank signals can be similarly treated.

In this report, we will require that the signal processing be unbiased. This restriction fixes the processing of the signal so that the signal may effectively be dropped from further consideration. The problem to be treated now is one of noise prediction only;  $c-1$  of the channels are used to predict the noise on the signal channel.



---

After reducing the signal extraction to noise prediction, many structures for predicting the noise on the principal channel are possible. This report will use the following linear estimation technique. Each of the noise-only c-seismometer time functions recorded during a time interval of  $T$  sec is individually Fourier-transformed at frequency points  $1/T$  apart. At any particular frequency, each channel will measure a single complex number representing the amplitude/phase of the noise on that channel at that frequency. The complex amplitude on the principal channel at that frequency can be estimated by using a linear combination of the  $c-1$  complex amplitudes. The complete Fourier transform of the principal channel noise can be estimated if this procedure is repeated for each frequency in the Fourier transforms. The predicted noise on the principal channel is then obtained by inverse Fourier transformation.

At each frequency, the  $c-1$  complex amplitudes are combined linearly to estimate the complex amplitude of the principal channel. This weighting of the channels is a complex filtering of the channels. The minimum variance filter weights (for the particular frequency) would be the Wiener filter weights based on the channel noise covariance (power spectra matrix) at that frequency. Typically, the noise covariance is different for different frequencies; hence, different complex weights are required for each frequency. Since the problem is identical at each frequency, the noise prediction problem will be treated at one particular frequency, which is understood to be any one of the frequencies of the Fourier transform.

In general, the noise covariance or power spectra matrix (at the understood frequency) is unknown and must be estimated in order to design the complex filter. By taking many  $T$ -sec time intervals, a series of samples is obtained which can be used to estimate the power spectra matrix, a  $c \times c$  Hermitian matrix, from which a filter can be designed. Generally, this estimated filter will not be the optimal Wiener filter but, hopefully, it will have



---

a mean absolute square error in predicting the complex amplitude of the principal channel noise reasonably close to the optimal Wiener filter performance. The ratio of the performance of the estimated filter to that of the Wiener filter indicates exactly how the estimated filter compares to the optimal filter. If the ratio is near unity, the estimated filter is almost as good as the Wiener filter and is satisfactory. If the ratio is much greater than unity, considerable improvement in the performance is possible by using more samples to design the estimated filter.

The relative performance (ratio) of the estimated filter is a random variable dependent on the noise samples. The probability density of the ratio, however, is independent of the actual noise covariance if the complex amplitudes of the noise are jointly zero-mean Gaussian. This fact is a fortuitous result for filter design. Generally, the noise power spectra matrix is unknown; this is the reason for estimating the spectra in the first place. Yet, the statistics of the performance ratio are calculable because the ratio is independent of the unknown noise covariance. By knowing the statistical properties of the performance ratio, the designer knows before any data are taken how well (relative to the optimal Wiener performance) the estimated filter is likely to perform. Before any experiment is conducted, the relative expected performance can be determined; the designer has prior knowledge as to whether the experiment will yield statistically accurate results. As an example of the type of information available, an estimated filter for a 5-channel array based on 31 independent samples will perform (with 90-percent confidence) within 1 db of the optimal (but unknown) Wiener filter. The properties of this performance ratio are derived and developed in the text.

A quantity which is available after the experiment is the regression error or prediction error of the estimated filter on the design samples. This quantity is an estimate of the performance of the estimated filter. The ratio of regression error to optimal Wiener error is also statistically independent of the noise power spectra matrix; thus, the statistical properties



---

of this regression-error ratio permit the designer to determine how well the regression error is likely to estimate the prediction error of the estimated filter. Generally, the regression error has some false gain; the properties of the regression-error ratio indicate the likely spread of the false gain.

Section II formulates the mathematical array prediction problem. Section III proves the invariance of the relative performance ratio and the regression-error ratio with respect to the actual noise covariance matrix. Section IV states the statistical properties derived in Appendix A, Section V gives some applications of the results to a filter design problem, and Section VI presents a matrix generalization of the invariance theorem.



---

## SECTION II

### MEAN-SQUARE ESTIMATION FORMULATION

The source samples  $\underline{x}_i$  (a single bar under a variable indicating a column vector, and a double bar indicating a square matrix) are assumed to be complex Gaussian random vectors satisfying the following definition, where  $t$  denotes conjugate transpose and  $|\underline{\Sigma}|$  is the determinant of the matrix  $\underline{\Sigma}$ .

A complex Gaussian random vector  $\underline{x}$  of  $c$  dimensions has a probability density (assuming zero mean) of

$$p \left[ \underline{x}^{(1)}, \underline{x}^{(2)}, \dots, \underline{x}^{(c)} \right] = p(\underline{x}) = \frac{1}{\pi^c |\underline{\Sigma}|} e^{-\underline{x}^t \underline{\Sigma}^{-1} \underline{x}} \quad (2-1)$$

where

$$\underline{\Sigma} = E \left[ \underline{x} \underline{x}^t \right] \quad (2-2)$$

Goodman<sup>1</sup> explains the relationship between the  $c$ -dimensional complex vector just defined and a  $2c$ -dimensional real Gaussian random vector.

Consider the problem of designing a linear filter  $\underline{f}$  (dimension  $c-1$ ) which predicts or estimates  $\underline{x}^{(1)}$  (the first component of  $\underline{x}$ ) from a linear combination of the other  $c-1$  variables  $\underline{x}^{(2)}, \dots, \underline{x}^{(c)}$ . The error in estimating  $\underline{x}^{(1)}$  is

$$\underline{e} = \underline{x}^{(1)} - \hat{\underline{x}}^{(1)} = \left[ 1, -\underline{f}^t \right] \underline{x}$$



---

In weighting the elements of  $\underline{x}$ , the complex conjugates of the elements of  $\underline{f}$  are used. However, this difficulty is more than offset by the notational simplification of the definitions used. The mean-square error  $\sigma^2(\underline{f})$  of any filter  $\underline{f}$  can be written

$$E[|\underline{e}|^2] = \sigma^2(\underline{f}) = \begin{bmatrix} 1, -\underline{f}^t \end{bmatrix} \underline{\Sigma} \begin{bmatrix} 1 \\ -\underline{f} \end{bmatrix} = \text{performance of filter } \underline{f} \quad (2-3)$$

The linear filter  $\underline{f}_0$  which produces the minimum mean-square error  $\sigma_0^2$  is the Wiener filter given by

$$\underline{\Sigma} \begin{bmatrix} 1 \\ -\underline{f}_0 \end{bmatrix} = \begin{bmatrix} \sigma_0^2 \\ 0 \end{bmatrix} \quad (2-4)$$

Assuming that  $\underline{\Sigma}$  is unknown but that  $n$  independent samples  $(\underline{x}_1, \dots, \underline{x}_n)$  are available, an estimate of the covariance matrix can be obtained as the sample covariance  $\hat{\underline{\Sigma}}$ :

$$\hat{\underline{\Sigma}} = \frac{1}{n} \sum_{k=1}^n \underline{x}_k \underline{x}_k^t \quad (2-5)$$



---

An estimate  $\hat{f}$  of the true optimum filter  $f_o$  is obtained as the solution of

$$\hat{\Sigma} = \begin{bmatrix} 1 \\ -\hat{f} \end{bmatrix} = \begin{bmatrix} \hat{\sigma}^2 \\ 0 \end{bmatrix} \quad (2-6)$$

where  $\hat{\sigma}^2$  is the regression or prediction error of the estimated filter operating on the  $n$  design points; i. e., the quantity  $\hat{\sigma}^2$  is the estimated performance of the estimated filter.

The true performance of the estimated filter  $\hat{f}$  is obtained by inserting  $\hat{f}$  from Equation (2-6) into Equation (2-3). This performance is defined as

$$\sigma^2 = \sigma^2(\hat{f}) = \begin{bmatrix} 1, -\hat{f}^t \end{bmatrix} \hat{\Sigma} \begin{bmatrix} 1 \\ -\hat{f} \end{bmatrix} \quad (2-7)$$

The true estimated-filter performance  $\sigma^2$  is a real positive random variable dependent on the  $n$  sample points and the covariance  $\hat{\Sigma}$ . Since the estimated filter cannot perform better than the optimal,  $\sigma^2$  is greater than or equal to  $\sigma_o^2$ . The ratio of  $\sigma^2$  to  $\sigma_o^2$  is a useful quantity for the filter designer. This ratio is defined as

$$\alpha \equiv \frac{\sigma^2}{\sigma_o^2} \geq 1 \quad (2-8)$$



---

If  $\alpha$  is close to unity, the designer knows that the estimated filter is nearly as good as the optimal and, therefore, is a well-designed filter. On the other hand, if  $\alpha$  is much greater than unity, the implication is that the estimated-filter performance can be significantly improved by using more samples in its design. Actually,  $\alpha$  can never be measured or calculated if the actual noise covariance matrix is unknown. However, surprisingly enough, for the Gaussian problem, the statistical properties of  $\alpha$  are independent of the actual noise covariance matrix  $\Sigma$ . In fact, the probability distribution of  $\alpha$  depends only on the number of channels and the number of samples used to design the filter. Since  $\alpha$  measures how near the estimated-filter performance is to the optimum and since the statistics of  $\alpha$  are independent of a particular noise covariance matrix, the filter designer can determine in advance the amount of data required to obtain a well-designed filter for a c-channel problem. For example, if the designer has enough data so that  $\alpha$  is less than 1.05 with a probability of 0.99, he knows that the estimated-filter performance will almost certainly be within 5 percent of the unknown optimal performance.

Another important random variable is the ratio of the regression error to the optimal Wiener error. This ratio is defined as  $\beta$ , where

$$\beta \equiv \frac{\hat{\sigma}_e^2}{\sigma_e^2} \geq 0 \quad (2-9)$$

Again, this ratio is independent of the noise covariance  $\Sigma$ . In designing the estimated filter, the regression error can be calculated. This quantity is the prediction error of the estimated filter on the design points and is an estimate of the estimated filter's mean-square error. The relative regression error  $\beta$  indicates how close the regression error is to the optimal Wiener error. Most of the time,  $\beta < 1$ , indicating a false gain, but there is the possibility that  $\beta \geq 1$ .



---

Since  $\beta$  and  $\alpha$  are also statistically independent of each other, their ratio,

$$\frac{\beta}{\alpha} = \frac{\hat{\sigma}^2}{\sigma^2}$$

also is independent of  $\Sigma$  and gives an intuitive idea of the relation between the regression error and the true error of the estimated filter.

In summary, the quantity  $\alpha$  is useful in determining how close to optimum an estimated filter can be expected to perform. The quantity  $\beta$  indicates the reliability of the regression error as a measure of  $\sigma_o^2$  and  $\sigma^2$ .

**BLANK PAGE**



---

### SECTION III

#### INVARIANCE PROPERTIES OF $\alpha$ AND $\beta$

Expressions for  $\alpha$  and  $\beta$ , which are independent of the noise covariance  $\underline{\Sigma}$ , are obtained. The quantities  $\sigma^2$  and  $\hat{\sigma}^2$  are linearly dependent only on  $\sigma_o^2$  and have weights  $\alpha$  and  $\beta$ , respectively.

To simplify the expression for  $\sigma^2$ , let

$$\underline{\Sigma} = \underline{A} \underline{A}^t \quad (3-1)$$

where  $\underline{A}$  is arbitrary except that  $a_{11} \neq 0$  and the remainder of the first column of  $\underline{A}$  is 0; e.g.,  $\underline{A}$  could be chosen upper triangular. There is no unique  $\underline{A}$  satisfying these conditions.  $\underline{A}$  is nonsingular because  $\underline{\Sigma}$  is assumed non-singular.

The Gaussian random vector  $\underline{x}$  can be visualized as being derived from another Gaussian random vector  $\underline{\xi}$ , with

$$\underline{x} = \underline{A} \underline{\xi} \quad (3-2)$$

and

$$E[\underline{\xi} \underline{\xi}^t] = \underline{I} = \text{identity} \quad (3-3)$$

It follows that, if

$$\hat{\underline{I}} = \frac{1}{n} \sum_{k=1}^n \underline{\xi}_k \underline{\xi}_k^t \quad (3-4)$$



then, from Equation (2-5),

$$\hat{\underline{\Sigma}} = \underline{\underline{A}} \hat{\underline{\underline{I}}} \underline{\underline{A}}^t \quad (3-5)$$

The expression (2-6) for the estimated filter  $\hat{\underline{f}}$  becomes

$$\underline{\underline{A}} \hat{\underline{\underline{I}}} \underline{\underline{A}}^t \begin{bmatrix} 1 \\ -\hat{f} \end{bmatrix} = \begin{bmatrix} \hat{\sigma}^2 \\ 0 \end{bmatrix} \quad (3-6)$$

Premultiplying by  $\frac{1}{a_{11}} \underline{\underline{A}}^{-1}$  in Equation (3-6) and grouping the factors to define the complex  $(c-1)$  dimensional  $\hat{\Gamma}$  yields

$$\hat{\underline{\underline{I}}} \begin{bmatrix} 1 \\ -\hat{\Gamma} \end{bmatrix} = \hat{\underline{\underline{I}}} \left( \frac{1}{a_{11}} \underline{\underline{A}}^{-1} \underline{\underline{A}}^t \begin{bmatrix} 1 \\ -\hat{f} \end{bmatrix} \right) = \begin{bmatrix} \frac{\hat{\sigma}^2}{|a_{11}|^2} \\ 0 \end{bmatrix} \quad (3-7)$$

Observe that Equation (3-7) is a canonical filter equation for a different problem, viz., estimating  $\xi^{(1)}$  from the other  $c-1$  variables  $\xi^{(2)}, \dots, \xi^{(c)}$ . Equation (3-7) is in the canonical form due to the particular choice of the form of  $\underline{\underline{A}}$ .



Since the  $\xi^{(i)}$ 's are uncorrelated, the optimal filter  $\underline{\Gamma}_o$  for this different problem is

$$\underline{\Gamma}_o = \underline{0} \quad (3-8)$$

Using Equation (3-7) as the definition of  $\hat{\underline{\Gamma}}$ , the true performance  $\sigma^2$  of the estimated filter  $\hat{\underline{f}}$  in Equation (2-7) is related to  $\hat{\underline{\Gamma}}$  by

$$\sigma^2 = |a_{11}|^2 \left( 1 + \hat{\underline{\Gamma}}^t \hat{\underline{\Gamma}} \right) \quad (3-9)$$

The performance of the optimal filter  $\underline{f}_o$  is

$$\sigma_o^2 = |a_{11}|^2 \left( 1 + \underline{\Gamma}_o^t \underline{\Gamma}_o \right) = |a_{11}|^2 \quad (3-10)$$

so that

$$\alpha = \frac{\sigma^2}{\sigma_o^2} = 1 + \hat{\underline{\Gamma}}^t \hat{\underline{\Gamma}} \quad (3-11)$$

Using Equation (2-9), Equation (3-7) becomes

$$\hat{\underline{I}} = \begin{bmatrix} 1 \\ \hat{\underline{\Gamma}} \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \end{bmatrix} \quad (3-12)$$

which is independent of  $\underline{\Sigma}$ , implying that  $\alpha$  and  $\beta$  do not depend on the original source covariance. The fact that  $|a_{11}|^2 = \sigma_o^2$  is also obvious from the definition of  $\underline{A}$  in Equation (3-1).



---

In this section, the estimation problem with covariance  $\underline{\Sigma}$  has been transformed to the same problem with covariance  $\underline{I}$ . The problem statistics remain the same only because the transformation is linear and the noise is Gaussian. If the noise were non-Gaussian, the same transformation would yield the  $\xi^{(i)}$ 's with covariance  $\underline{I}$ , but the statistics of the  $\xi_i^{(i)}$ 's would not necessarily be tractable. Under the Gaussian assumption, all covariances  $\underline{\Sigma}$  have identical statistics for  $\alpha$  and  $\beta$ ; hence, the densities of  $\alpha$  and  $\beta$  are independent of  $\underline{\Sigma}$ .



---

## SECTION IV

### PROBABILITY DENSITIES OF $\alpha$ AND $\beta$

After the original problem with covariance  $\underline{\Sigma}$  is reduced to the same problem with covariance  $\underline{\underline{\Sigma}}$ , the joint probability density of  $\alpha$  and  $\beta$  follows (Appendix A) from the results in Goodman.<sup>1</sup>  $\alpha$  and  $\beta$  are independent, with densities given in Table IV-1. The situation for  $n < c$  involves a degenerate  $\hat{\underline{\Sigma}}$  and is not treated here.

The densities for  $\alpha$  and  $\beta$  are easily related to  $\chi^2$  random variables. The probability density of  $2n\beta$  is  $\chi^2$  with  $2(n-c+1)$  degrees of freedom. If  $\chi_1^2$  and  $\chi_2^2$  are independent  $\chi^2$  variables with  $v_1 = 2(c-1)$  and  $v_2 = 2(n-c+2)$  degrees of freedom, the random variable

$$1 + \frac{\chi_1^2}{\chi_2^2}$$

has the same density as  $\alpha$  (Table IV-1). The quantity

$$\frac{\chi_1^2/v_1}{\chi_2^2/v_2} = \frac{v_2}{v_1} (\alpha - 1)$$

is F-distributed. The quantity  $1/\alpha$  is  $\beta$ -distributed.

These distributions are common and tabulated in the NBS Handbook of Mathematical Functions.<sup>2</sup>



Table IV-1  
PROBABILITY DENSITIES OF  $\alpha$  AND  $\beta$

$\alpha = \text{ratio of performance}$	$\frac{\text{estimated-filter error}}{\text{optimal-filter error}} = \frac{\sigma^2}{\sigma_o^2}$
Density	$\frac{n!}{(n-c+1)! (c-2)!} \frac{(\alpha-1)^{c-2}}{\alpha^{n+1}}$
Region	$\alpha \geq 1, n \geq c \geq 2$
Mean	$\frac{n}{n-c+1}$
Variance	$\frac{n(c-1)}{(n-c+1)^2 (n-c)}$
$\beta = \text{ratio of}$	$\frac{\text{regression error of estimated filter}}{\text{optimal-filter error}} = \frac{\hat{\sigma}^2}{\sigma_o^2}$
Density	$\frac{n^{n-c+1} \beta^{n-c} e^{-n\beta}}{(n-c)!}$
Region	$\beta \geq 0, n \geq c \geq 2$
Mean	$\frac{n-c+1}{n}$
Variance	$\frac{n-c+1}{n^2}$

If the random vector  $\underline{x}$  is actually real Gaussian with covariance  $\Sigma$ , the densities for  $\alpha$  and  $\beta$  are only slightly altered. In the density for  $\alpha$  (Table IV-1), if  $n+1$  is replaced by  $\frac{n+1}{2}$  and if  $c-1$  is replaced by  $\frac{c-1}{2}$ , the resulting density is the density for  $\alpha$  if  $\underline{x}$  is a real Gaussian vector. Similarly, if  $c-1$  is replaced by  $\frac{c-1}{2}$  and  $n$  by  $\frac{n}{2}$  in  $p(\beta)$ , the new density is correct for  $\beta$  if  $\underline{x}$  is real. These changes follow from a direct evaluation of  $p(\alpha, \beta)$  for  $\underline{x}$  real and are not necessarily obvious.

**BLANK PAGE**



---

## SECTION V APPLICATIONS

The follow is an example to demonstrate how the results in Section IV are applied.

An array of 11 horizontal seismometers is used to predict the noise on the principal vertical seismometer. The vertical signal incident on the array may be ignored in the analysis; the array predicts the noise and not the signal on the principal seismometer. The noise crosspower values between the 12 channels are unknown and must be estimated from measured Fourier transform data; 80 noise samples per frequency are available.

### A. PROPERTIES OF $\alpha$

The performance ratio  $\alpha$  represents the mean-square error of the estimated filter relative to that of the Wiener filter. If  $\alpha$  is near unity, the estimated filter performs almost as well as the optimal filter. Statistically, the value of  $\alpha$ , obtained from a particular sample of  $n$  design points, depends on how well these  $n$  points represent the true covariance matrix. The properties of  $\alpha$  are now determined for the above seismic example.

#### 1. Average Behavior

A filter designed for the above example performs on the average,

$$E[\alpha] = \frac{n}{n - c + 1} = 1.159 = 0.64 \text{ db}$$

times the optimal minimum mean-square error.



The standard deviation gives some idea of the possible spread to expect. A random variable (with a reasonably high probability) can be said to be within two standard deviations of its mean. For this example,

$$\text{standard deviation} = \sqrt{\frac{n(c-1)}{(n-c+1)^2(n-c)}} = 0.052$$

Therefore, the performance of the estimated filter is (with high probability) between 1.055 and 1.263 times the optimal error. More precise statements can be made by considering the exact density.

## 2. Distribution Function

The integrated probability density or distribution of  $\alpha$  indicates the manner in which  $\alpha$  can be spread. Figure V-1 shows the plot of

$$\begin{aligned} P(\alpha_o) &= \int_1^{\alpha_o} p(\alpha) d\alpha \\ &= P[\alpha \leq \alpha_o] \end{aligned}$$

for  $n = 80$  and  $c = 12$ .

The exact probability of being within two standard deviations can be calculated (using Figure V-1) as

$$P[1.055 < \alpha \leq 1.263] = P(1.263) - P(1.055) = 0.96$$

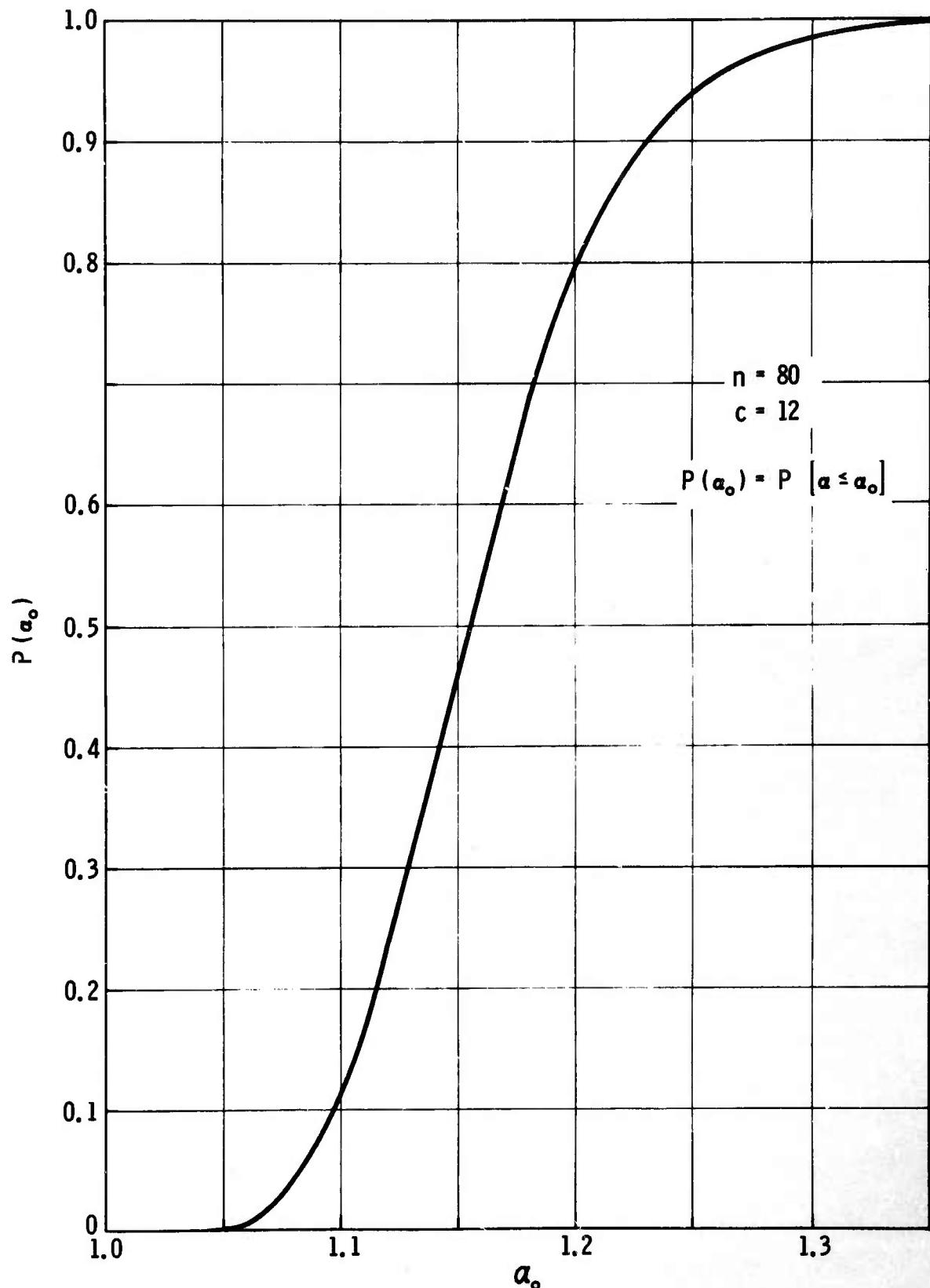


Figure V-1. Cumulative Distribution of  $\alpha$  for 12-Channel Array and 80 Samples



---

Other useful information can be read off the distribution function. For example, the estimated-filter performance will be no more than 1.2 times the optimal 80 percent of the time. The estimated filter will be better than 1.1 times the optimal filter only 10 percent of the time. Seldom will the estimated-filter performance be 1.05 times or less than the optimal.

### 3. Density Function

Figure V-2 plots the probability density of  $\alpha$  as a function of the number of samples used to estimate the covariance matrix. The density is quite broad for  $n = 20$ , indicating that  $\alpha$  probably will be much greater than unity. Thus, the estimated-filter performance is likely to be several times the Wiener filter performance. As  $n$  increases, the density becomes more peaked and approaches unity, indicating the improvement gained by increasing the number of samples used to design the filter.

## B. CONFIDENCE-LEVEL PLOTS

Another presentation of the properties of  $\alpha$  is made in Figures V-3 through V-6. The plots are of the number of samples necessary if estimated-filter performance is to be within 2 db, 1 db, 0.5 db, or 0.1 db of the Wiener filter performance for confidence levels of 50 percent, 80 percent, 90 percent, and 99 percent. If  $c = 12$  (as in the example),  $n = 80$  samples is more than enough for 90-percent confidence of being within 1 db of optimal (Figure V-5). For a given confidence level and relative performance (in decibels), these plots specify the number of samples needed.

## C. APPROXIMATIONS

Asymptotically, the tail of the  $\alpha$  density approaches the tail of the normal density. Empirically, in levels less than 1 db and  $n > 5c$ , there is very little error in calculating Figure V-3 by assuming that  $\alpha$  is normally distributed with mean and variance as given in Table IV-1.

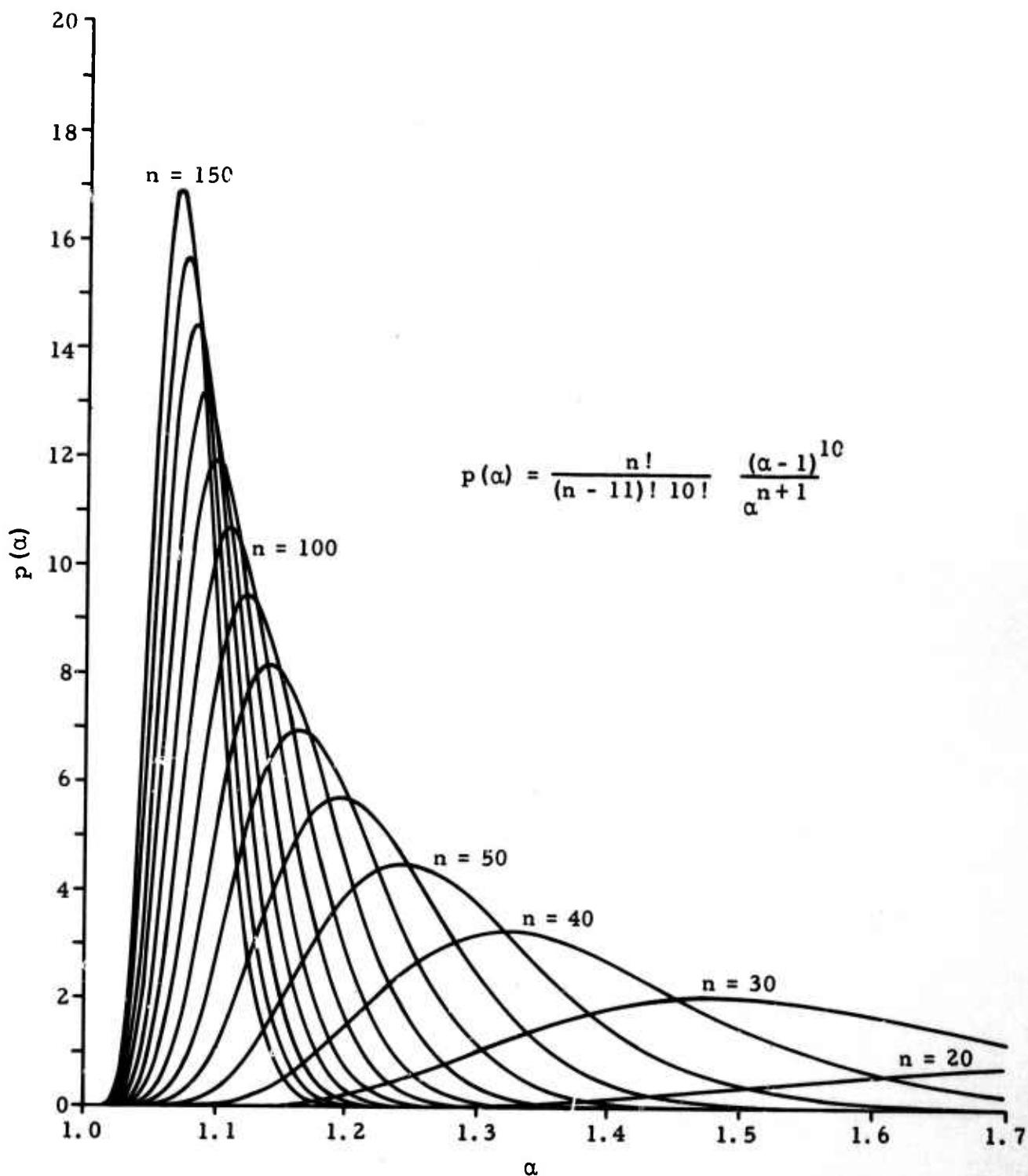


Figure V-2. Probability Densities of  $\alpha$  for  $c = 1/2$  Channels

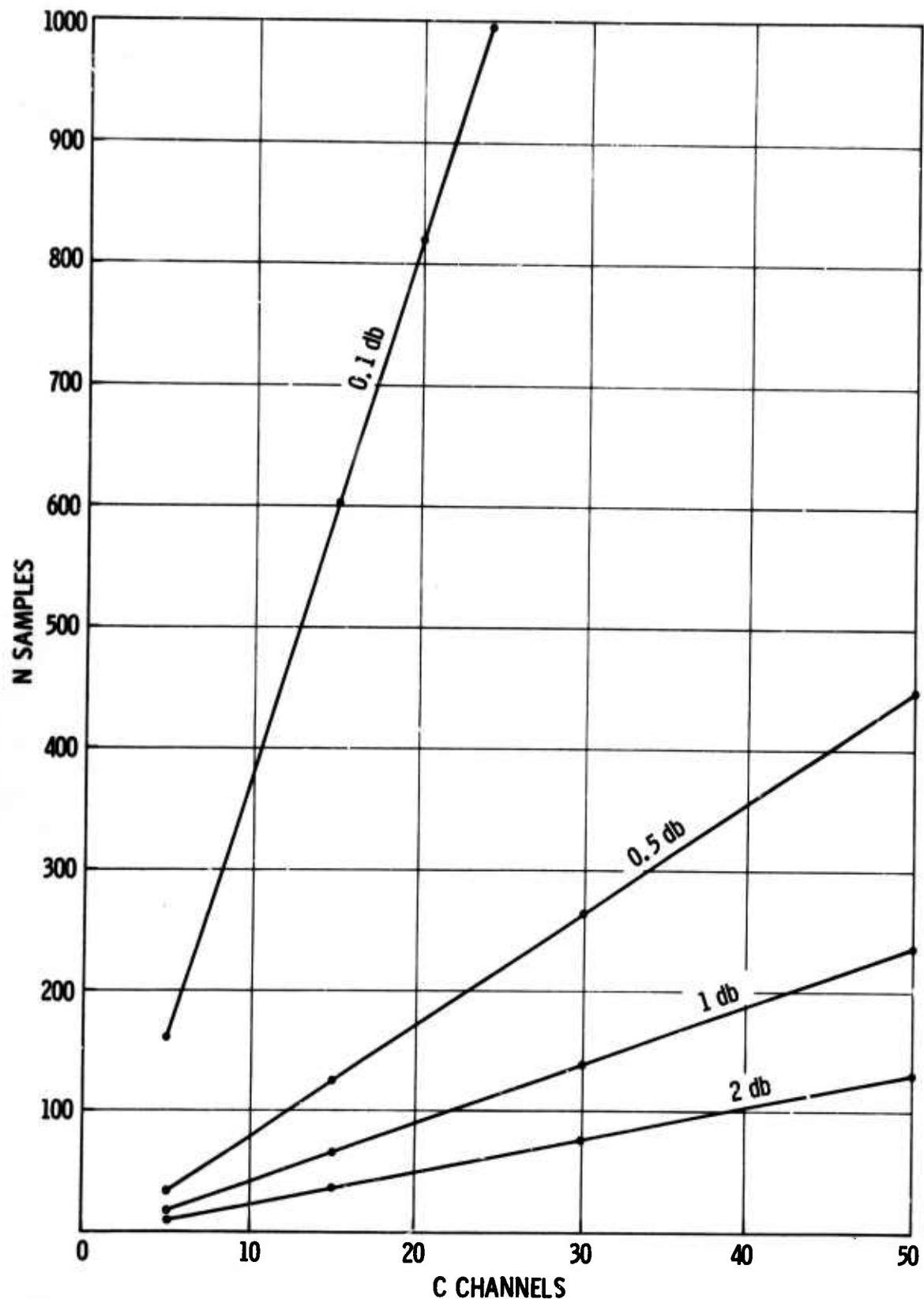


Figure V-3. 50-Percent Confidence Plot

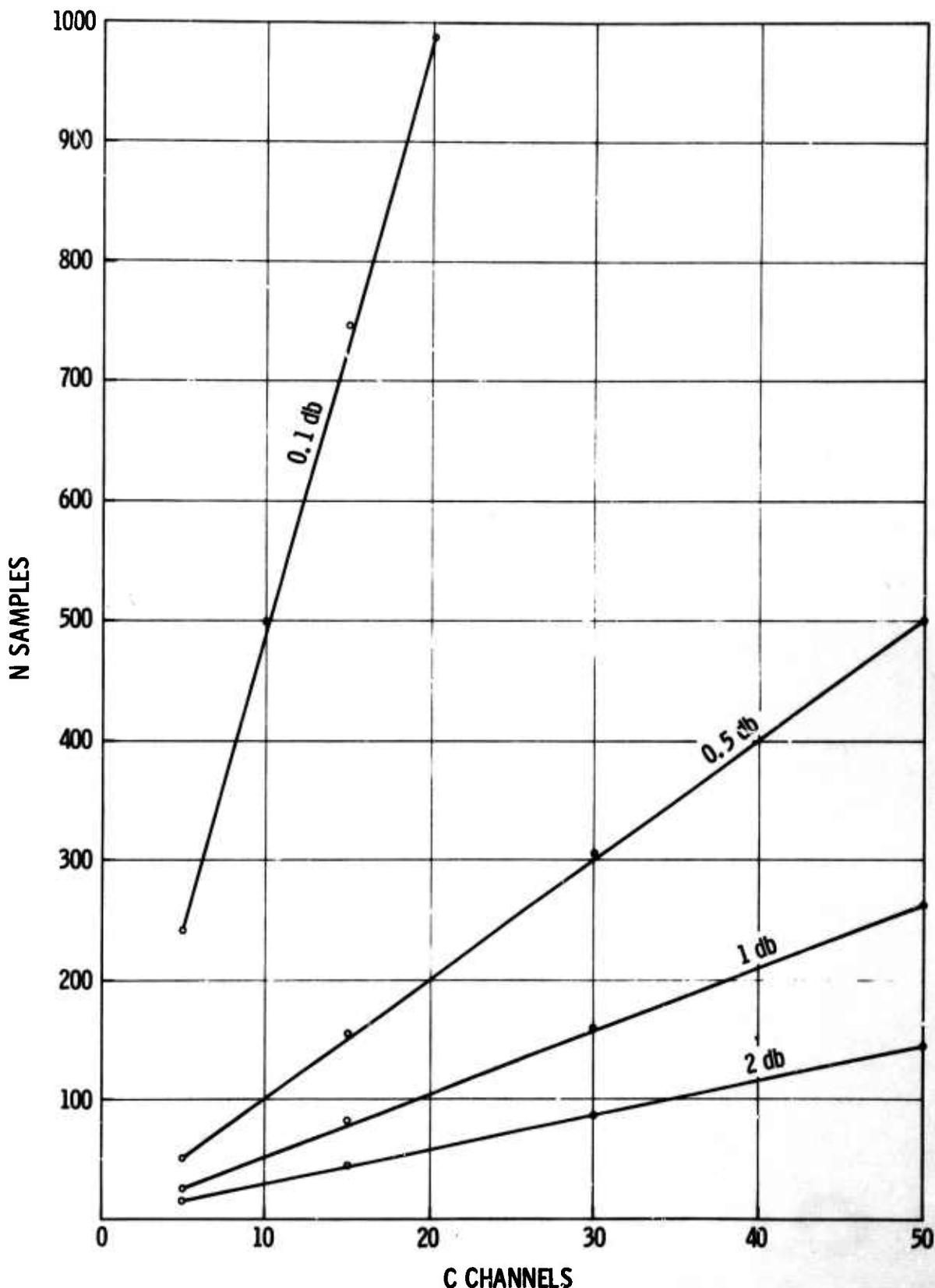


Figure V-4. 80-Percent Confidence Plot

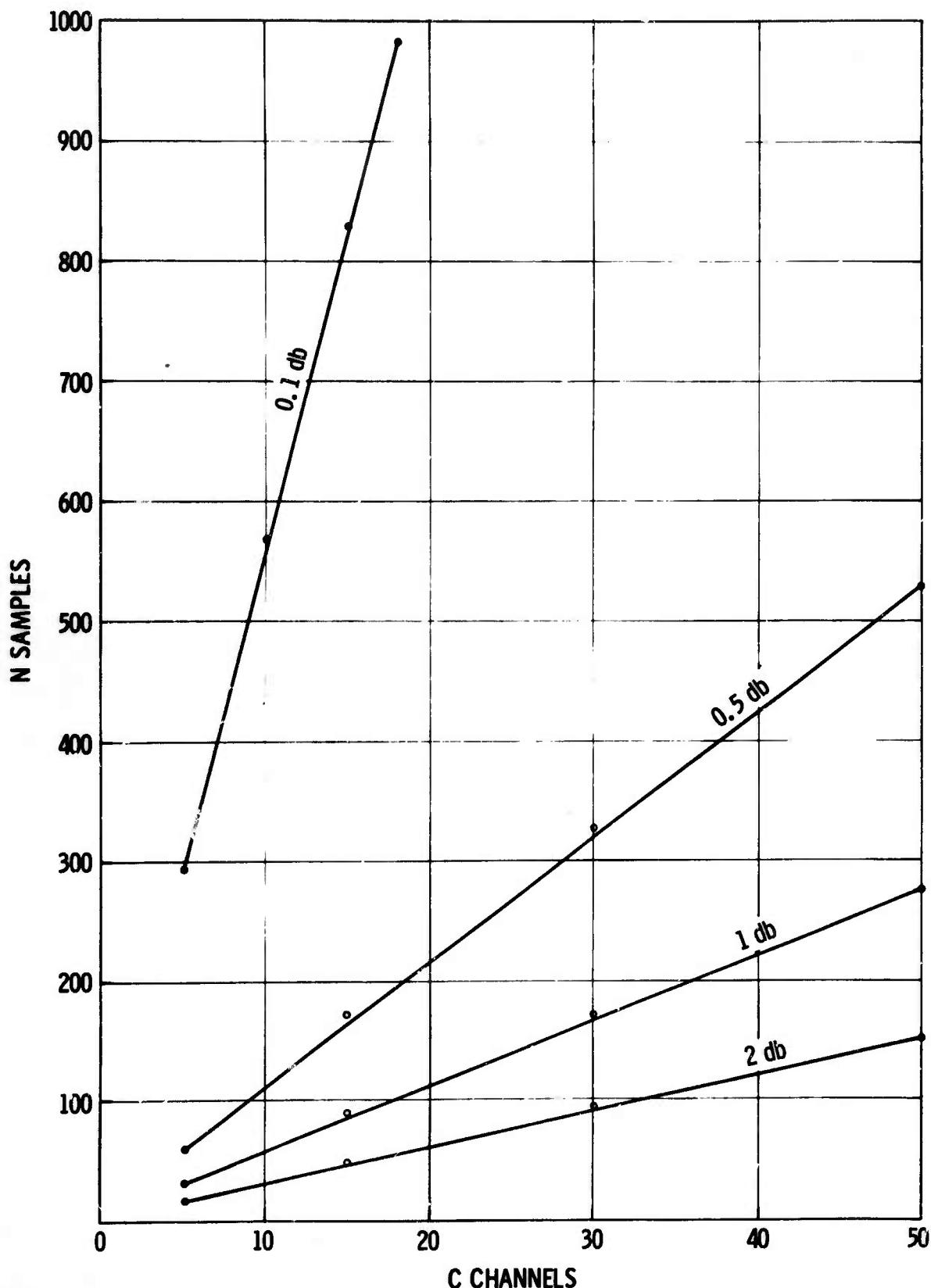


Figure V-5. 90-Percent Confidence Plot

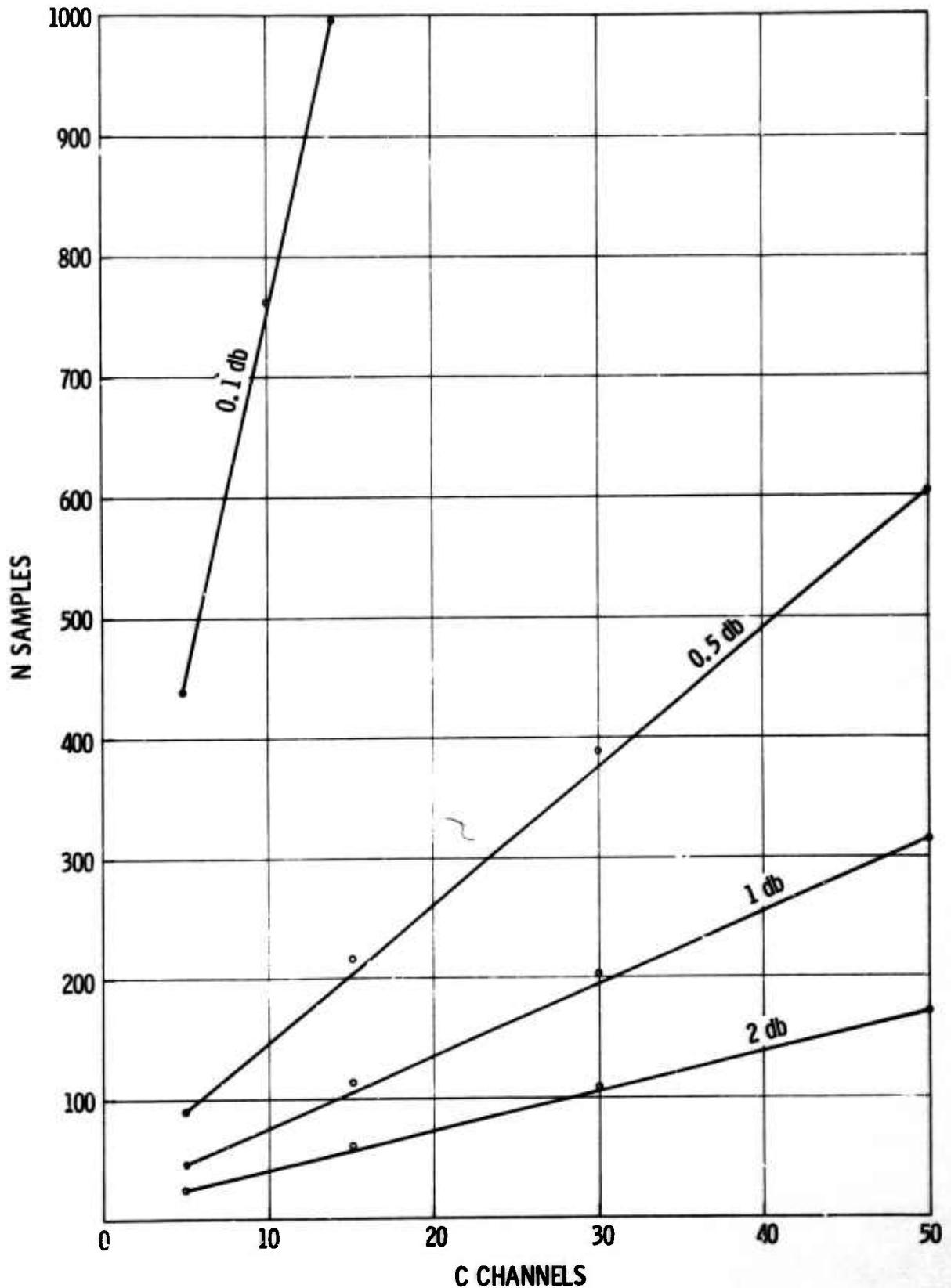


Figure V-6. 99-Percent Confidence Plot



---

The 90-percent level corresponds to no more than 1.282 standard deviations to the right of the mean. Thus, for  $n > 5c$ , the 1-db line is approximately

$$1.26 = \frac{n}{n - c + 1} + 1.282 \sqrt{\frac{n(c - 1)}{(n - c + 1)^2 (n - c)}}$$

This type of approximation can be used for most cases which are not given explicitly in the figures.

#### D. PROPERTIES OF $\beta$

The relative performance of the estimated filter is indicated by  $\alpha$  but cannot be measured. The regression error, on the other hand, is calculated during filter design and relates to the actual performance of the estimated filter. As previously stated, the regression error (or false gain) is the prediction error of the estimated filter operating on the  $n$  design points.  $\beta$  is the ratio of the regression error to the true optimal error.

If  $\beta < 1$ , the regression error is less than the optimal error and represents false gain; however, if  $\beta > 1$ , the regression error is greater than the optimal-filter error. Note that only the regression error ( $\beta \sigma_o^2$ ) is measurable;  $\beta$  is not.

##### 1. Average Behavior

The average  $\beta$  obtained in a filter design is

$$E[\beta] = \frac{n - c + 1}{n} = 1 - \frac{c - 1}{n}$$



---

This is less than 1.0, indicating an expected false gain. An unbiased estimate of the optimal minimum mean-square error is

$$\text{unbiased estimate of } \sigma_o^2 = \frac{n}{n - c + 1} \hat{\sigma}^2$$

This scales up the regression error by its average false gain.

The actual performance of the estimated filter is greater than  $\sigma_o^2$  by an average amount

$$E[\alpha] = \frac{n}{n - c + 1}$$

Thus, the result will approximate the performance of the estimated filter if the regression error is scaled by

$$\frac{E[\alpha]}{E[\beta]} = \frac{n^2}{(n - c + 1)^2}$$

Roughly, the estimated-filter performance is

$$\frac{n^2}{(n - c + 1)^2}$$

times the measured regression error. For 12 channels and 80 samples, the true performance of the estimated filter is (on the average) 1.35 times the actual measured regression error.



## 2. Density Function

Figure V-7 plots the densities of  $\beta$  corresponding to the densities of  $\alpha$  in Figure V-2. For  $n = 20$ ,  $\beta$  is likely to be much less than unity, indicating that the regression error is likely to be much less than the Wiener filter error. Thus, the false gain is likely to be great. As  $n$  increases, the density of  $\beta$  becomes more peaked near unity.

The designer can look at these plots and decide how many samples are needed to obtain the estimated filter. Figure V-2 indicates the spread in relative performance to be expected, and Figure V-7 indicates the spread in regression error (false gain) to be expected.

As stated earlier,  $\alpha$  and  $\beta$  are independent. Even if  $\alpha = 1$  (i.e., estimated filter = optimal filter),  $\beta$  has the same density and is no better an estimate of  $\sigma_o^2$ . In fact, to estimate  $\sigma_o^2$  from the regression error with the optimal filter requires only  $c-1$  fewer samples (for equivalent estimates of  $\sigma_o^2$ ) than to estimate  $\sigma_o^2$  without the optimal filter. This is analogous to estimating a Gaussian variance with and without the mean. Without the mean, one additional sample is required for equivalent estimates of the variance.

## E. RESTRICTIONS AND EXTENSIONS

This report assumes throughout that  $n$  independent samples are available for the estimated filter. If the samples are actually correlated, the effective  $n$  will be less than the number of samples. Analysis for correlated samples is needed.

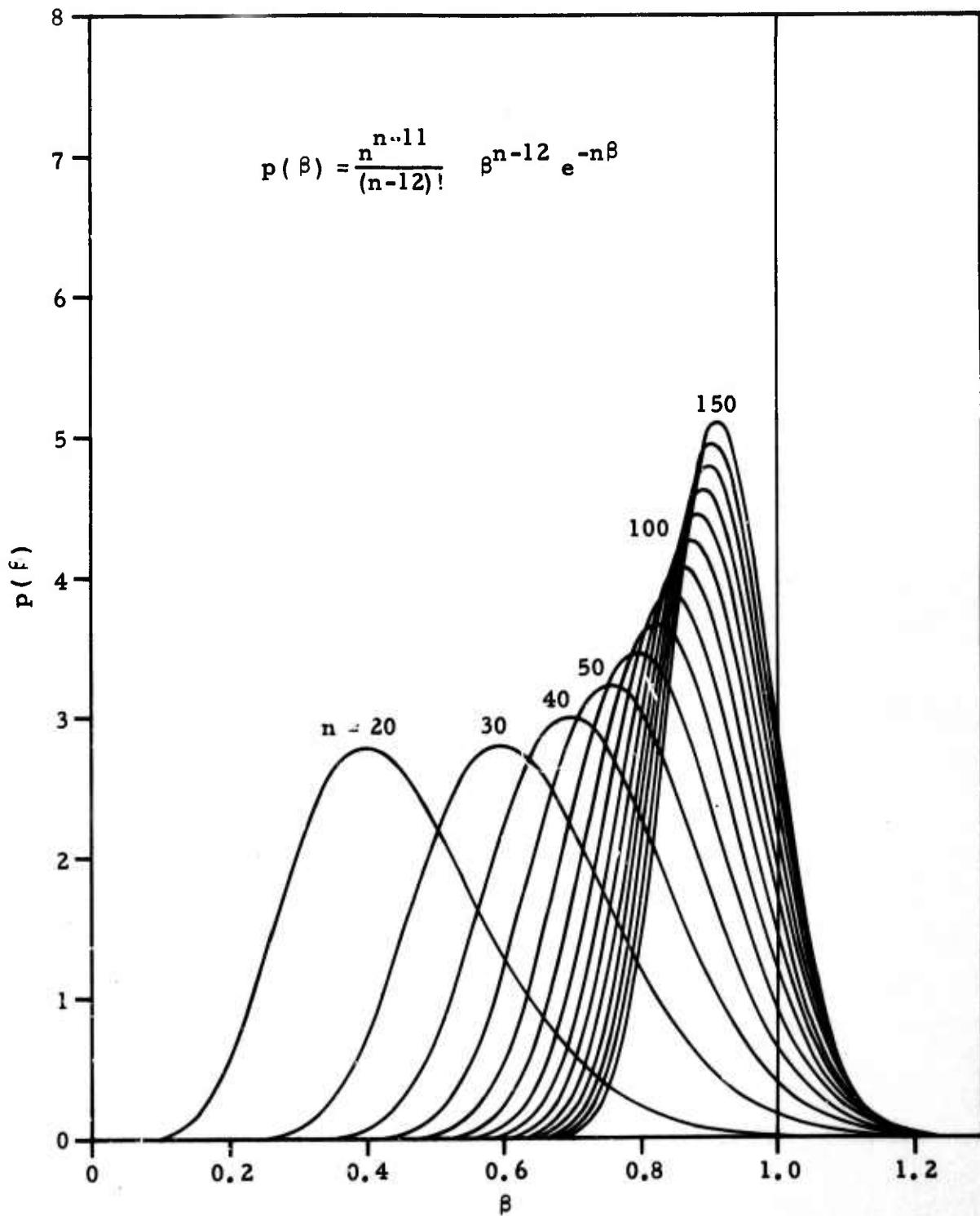


Figure V-7. Probability Densities of  $\beta$  for  $c = 12$  Channels

---

# BLANK PAGE



---

## SECTION VI

### GENERALIZED INVARIANCE PROPERTIES

This section discusses a generalization of the parameters  $\alpha$  and  $\beta$  to matrix parameter  $\underline{\underline{A}}$  and  $\underline{\underline{B}}$ . Consider a complex random variable  $\underline{\underline{x}}$  of dimension  $c$  with covariance matrix  $\underline{\underline{\Sigma}}_{\underline{\underline{x}}}$ . The covariance matrix  $\underline{\underline{\Sigma}}_{\underline{\underline{x}}}$  can be uniquely represented by  $\underline{\underline{\Sigma}}_{\underline{\underline{x}}} = \underline{\underline{L}}^T \underline{\underline{L}}$  where  $\underline{\underline{L}}$  is lower triangular (see Lemma 2 of Appendix A). If the representation is expressed in the form

$$\underline{\underline{\Sigma}} \underline{\underline{L}}^{-1} \underline{\underline{D}} = \underline{\underline{L}}^T \underline{\underline{D}} \quad (6-1)$$

where  $\underline{\underline{D}}$  is a diagonal matrix consisting of the diagonal elements of  $\underline{\underline{L}}$ , the form is seen to be a matrix generalization of Equation (2-4). The columns of  $\underline{\underline{L}}^{-1} \underline{\underline{D}}$  represent prediction-error filters of varying length; the covariance matrix of the filter errors is

$$E \underline{\underline{D}} (\underline{\underline{L}}^{-1})^T \underline{\underline{x}} \underline{\underline{x}}^T \underline{\underline{L}}^{-1} \underline{\underline{D}} = \underline{\underline{D}} \underline{\underline{D}} \quad (6-2)$$

An analogous generalization can be made for  $\hat{\underline{\underline{L}}}$ ,  $\hat{\underline{\underline{D}}}$  derived from the estimated covariance matrix  $\hat{\underline{\underline{\Sigma}}}_{\underline{\underline{x}}}$  of Equation (2-5).

The generalization of  $\alpha$  to  $\underline{\underline{A}}$  is now achieved by normalizing the covariance matrix of the errors of the estimated filters

$$\hat{\underline{\underline{D}}}_{\underline{\underline{x}}} (\hat{\underline{\underline{L}}}_{\underline{\underline{x}}}^{-1})^T \hat{\underline{\underline{\Sigma}}}_{\underline{\underline{x}}} \hat{\underline{\underline{L}}}_{\underline{\underline{x}}}^{-1} \hat{\underline{\underline{D}}}_{\underline{\underline{x}}}$$

to get

$$\underline{\underline{A}}_{\underline{\underline{x}}} = \underline{\underline{D}}_{\underline{\underline{x}}}^{-1} \hat{\underline{\underline{D}}}_{\underline{\underline{x}}} (\hat{\underline{\underline{L}}}_{\underline{\underline{x}}}^{-1})^T \hat{\underline{\underline{\Sigma}}}_{\underline{\underline{x}}} \hat{\underline{\underline{L}}}_{\underline{\underline{x}}}^{-1} \hat{\underline{\underline{D}}}_{\underline{\underline{x}}}^{-1} \quad (6-3)$$



Similarly,  $\underline{\underline{B}}$  generalizes to

$$\underline{\underline{B}}_{\underline{\underline{x}}} = \underline{\underline{D}}_{\underline{\underline{x}}}^{-1} \underline{\underline{D}}_{\underline{\underline{x}}} (\underline{\underline{L}}_{\underline{\underline{x}}}^{-1})^T \underline{\underline{\Sigma}}_{\underline{\underline{x}}} \underline{\underline{L}}_{\underline{\underline{x}}}^{-1} \underline{\underline{D}}_{\underline{\underline{x}}}^{-1} \quad (6-4)$$

The matrices  $\underline{\underline{A}}$  and  $\underline{\underline{B}}$  are now evaluated for the random variable  $\underline{\underline{y}} = (\underline{\underline{L}}_{\underline{\underline{x}}}^T)^{-1} \underline{\underline{x}}$  and the observed data  $\underline{\underline{y}}_i = (\underline{\underline{L}}_{\underline{\underline{x}}}^T)^{-1} \underline{\underline{x}}_i$ .

First, note that

$$\underline{\underline{\Sigma}}_{\underline{\underline{y}}} = (\underline{\underline{L}}_{\underline{\underline{x}}}^T)^{-1} \underline{\underline{\Sigma}}_{\underline{\underline{x}}} \underline{\underline{L}}_{\underline{\underline{x}}}^{-1} = \underline{\underline{I}}$$

and

(6-5)

$$\underline{\underline{\Sigma}}_{\underline{\underline{y}}} = (\underline{\underline{L}}_{\underline{\underline{x}}}^T)^{-1} \underline{\underline{\Sigma}}_{\underline{\underline{x}}} \underline{\underline{L}}_{\underline{\underline{x}}}^{-1} = (\underline{\underline{L}}_{\underline{\underline{x}}}^T)^{-1} \underline{\underline{L}}_{\underline{\underline{x}}}^T \underline{\underline{\Sigma}}_{\underline{\underline{x}}} \underline{\underline{L}}_{\underline{\underline{x}}}^{-1}$$

so that

$$\underline{\underline{L}}_{\underline{\underline{y}}} = \underline{\underline{L}}_{\underline{\underline{x}}} \underline{\underline{L}}_{\underline{\underline{x}}}^{-1}$$

and

(6-6)

$$\underline{\underline{D}}_{\underline{\underline{y}}} = \underline{\underline{D}}_{\underline{\underline{x}}} \underline{\underline{D}}_{\underline{\underline{x}}}^{-1}$$

These relationships are now used to express

$$\underline{\underline{A}}_{\underline{\underline{y}}} = \underline{\underline{D}}_{\underline{\underline{y}}} (\underline{\underline{L}}_{\underline{\underline{y}}}^{-1})^T (\underline{\underline{L}}_{\underline{\underline{y}}}^{-1})^{-1} \underline{\underline{D}}_{\underline{\underline{y}}} \quad (6-7)$$

in the form

$$\underline{\underline{A}}_{\underline{\underline{y}}} = \underline{\underline{D}}_{\underline{\underline{x}}}^{-1} \underline{\underline{D}}_{\underline{\underline{x}}} (\underline{\underline{L}}_{\underline{\underline{x}}}^{-1})^T \underline{\underline{L}}_{\underline{\underline{x}}}^T \underline{\underline{L}}_{\underline{\underline{x}}} (\underline{\underline{L}}_{\underline{\underline{x}}}^{-1})^{-1} \underline{\underline{D}}_{\underline{\underline{x}}} \underline{\underline{D}}_{\underline{\underline{x}}}^{-1} = \underline{\underline{A}}_{\underline{\underline{x}}} \quad (6-8)$$

Similarly,  $\underline{\underline{B}}_{\underline{\underline{y}}} = \underline{\underline{B}}_{\underline{\underline{x}}}$ , so the distribution of the matrices  $\underline{\underline{A}}$  and  $\underline{\underline{B}}$  is independent of the covariance of the vector for which they are defined.



---

SECTION VII  
REFERENCES

1. Goodman, N. R., 1963, Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction): Ann. of Math. Statistics, v. 34, p. 152-176.
2. National Bureau of Standards, 1964: Handbook of Mathematical Functions.



---

APPENDIX A  
DERIVATION OF PROBABILITY DENSITIES OF  $\hat{\Gamma}$ ,  $\alpha$ , AND  $\theta$



---

## APPENDIX A

### DERIVATION OF PROBABILITY DENSITIES OF $\hat{\underline{\Gamma}}$ , $\alpha$ , AND $\beta$

The text equations specifying the filter  $\hat{\underline{\Gamma}}$  (c-1 dimensions)  
and the ratios  $\alpha$  and  $\beta$  are

$$\hat{\underline{\Gamma}} = \begin{bmatrix} 1 \\ -\hat{\underline{\Gamma}} \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \end{bmatrix}$$

$$\alpha = 1 + \hat{\underline{\Gamma}}^t \hat{\underline{\Gamma}} \quad (A-1)$$

and

$$\hat{\underline{\Gamma}} = \text{sample covariance} = \frac{1}{n} \sum_{k=1}^n \xi_k \xi_k^t$$

where  $\xi_k$  are complex Gaussian random vectors (zero-mean, c-dimensional)  
with covariance

$$E[\xi_k \xi_j^t] = \underline{\Gamma}$$



Prior to evaluating  $p(\alpha, \beta)$ , the required results from Goodman<sup>1</sup> are given.

• Lemma 1

If  $\underline{x}_1, \dots, \underline{x}_n$  are independent samples from a complex Gaussian (zero-mean,  $c$ -dimensional for  $c \leq n$ ) source with a covariance

$$E \left[ \underline{x}_k \underline{x}_k^t \right] = \underline{\Sigma}$$

then,

$$B = \sum_{i=1}^n \underline{x}_i \underline{x}_i^t \quad (A-2)$$

has a probability density

$$\begin{aligned} p(\underline{B}) &= p(B_{11}, B_{12}, \dots, B_{1c}, B_{22}, B_{23}, \dots, B_{cc}) \\ &= \frac{|\underline{B}|^{n-c}}{F(n, c, \underline{\Sigma})} \exp \left[ -\text{Tr}(\underline{\Sigma}^{-1} \underline{B}) \right] \end{aligned} \quad (A-3)$$

where

$$F(n, c, \underline{\Sigma}) = \pi^{\frac{1}{2}c(c-1)} (n-1)! \dots (n-c)! |\underline{\Sigma}|^n \quad (A-4)$$

The density above is defined on positive semi-definite Hermitian matrices  $\underline{B}$ . The density is the complex Wishart distribution.



• Corollary

If  $\underline{\underline{A}}$  and  $\underline{\underline{B}}$  are (cxc) Hermitian and  $\underline{\underline{A}}^{-1}$  exists, then

$$\int \cdots \int d\underline{\underline{B}} |\underline{\underline{B}}|^{n-c} \exp \left[ -\text{Tr}(\underline{\underline{A}} \underline{\underline{B}}) \right] = F(n, c, \underline{\underline{A}}^{-1}) \quad (A-5)$$

where  $\underline{\underline{B}}$  is positive semi-definite and

$$d\underline{\underline{B}} = dB_{11} dB_{12} \cdots dB_{1c} dB_{22} \cdots dB_{cc}$$

• Lemma 2

If  $\underline{\underline{B}}$  is a (cxc) complex random matrix satisfying the Wishart distribution in Lemma 1, there is a unique lower triangular complex matrix  $\underline{\underline{L}}$  with positive real diagonal elements  $L_{11}, \dots, L_{cc}$ , so that

$$\underline{\underline{L}}^t \underline{\underline{L}} = \underline{\underline{B}}$$

and

$$p(\underline{\underline{L}}) = p(L_{11}, L_{21}, \dots, L_{c1}, L_{22}, \dots, L_{c2}, L_{33}, \dots, L_{cc}) \quad (A-6)$$

$$= \frac{2^c}{F(n, c, \underline{\underline{\Sigma}})} L_{cc}^{2n-1} L_{c-1, c-1}^{2n-3} \cdots L_{11}^{2n-2c+1} \exp \left[ -\text{Tr}(\underline{\underline{\Sigma}}^{-1} \underline{\underline{L}}^t \underline{\underline{L}}) \right]$$

where  $L_{kk}$  is real and positive,  $L_{jk}$  ( $j > k$ ) is complex, and  $L_{jk}$  ( $j < k$ ) = 0.



Lemma 1 permits the writing of the probability density of the matrix  $\underline{\underline{I}}^{\wedge}$ . The analysis proceeds by changing variables first to  $\underline{\underline{L}}$  (lower triangular matrix) and then to a set of variables including  $\underline{\underline{\Gamma}}^{\wedge}$ . Finally, by integrating out all unwanted variables,  $p(\underline{\underline{\Gamma}}^{\wedge}, \underline{\underline{\beta}})$  remains. The joint density  $p(\underline{\underline{\alpha}}, \underline{\underline{\beta}})$  follows easily by a final change of variables.

From Lemma 1,  $\underline{\underline{B}}$  is just  $n$  times the sample covariance. Therefore, defining

$$\underline{\underline{B}} = n \underline{\underline{I}}^{\wedge}$$

and

$$\underline{\underline{L}}^t \underline{\underline{L}} = \underline{\underline{B}}$$

then,

$$\frac{1}{n} \underline{\underline{L}}^t \underline{\underline{L}} = \begin{bmatrix} 1 \\ -\underline{\underline{\Gamma}}^{\wedge} \end{bmatrix} = \begin{bmatrix} \underline{\underline{\beta}} \\ 0 \end{bmatrix} \quad (A-7)$$

Since  $\underline{\underline{L}}$  is lower triangular,  $\underline{\underline{L}}^{-1}$  is also lower triangular. The definition of  $\underline{\underline{\Gamma}}^{\wedge}$ , then becomes

$$\underline{\underline{L}} = \begin{bmatrix} 1 \\ -\underline{\underline{\Gamma}}^{\wedge} \end{bmatrix} = \begin{bmatrix} \frac{n\underline{\underline{\beta}}}{L_{11}} \\ 0 \end{bmatrix} = \begin{bmatrix} L_{11} \\ 0 \end{bmatrix} \quad (A-8)$$

and, defining  $\underline{\underline{\beta}}$  in terms of  $\underline{\underline{L}}$ ,

$$\underline{\underline{\beta}} = \frac{L_{11}^2}{n} \quad (A-9)$$



To make the change of variables  $L_{21}, \dots, L_{c1}$  to  $\hat{\Gamma}^{(1)}, \dots, \hat{\Gamma}^{(c-1)}$  but to retain the other elements of  $\underline{\underline{L}}$ , we partition  $\underline{\underline{L}}$  as

$$\underline{\underline{L}} = \begin{bmatrix} L_{11} & \underline{0} \\ \underline{b} & \underline{\underline{M}} \end{bmatrix} \quad (A-10)$$

where

$$\underline{b} = \begin{bmatrix} L_{21} \\ \vdots \\ L_{c1} \end{bmatrix}$$

The partition in Equation (A-10) implies that the equation defining the change of variables is

$$\underline{b} - \underline{\underline{M}} \hat{\Gamma} = \underline{0} \quad (\text{c-1 dimension}) \quad (A-11)$$

The Jacobian of this transformation is

$$|\underline{\underline{M}}|^2 = L_{22}^2 L_{33}^2 \dots L_{cc}^2 = \frac{\partial (b_R^{(1)}, b_I^{(1)}, b_R^{(2)}, \dots, b_I^{(c-1)})}{\partial (\hat{\Gamma}_R^{(1)}, \hat{\Gamma}_I^{(1)}, \hat{\Gamma}_R^{(2)}, \dots, \hat{\Gamma}_I^{(c-1)})} \quad (A-12)$$



The density before the transformation (Lemma 2) is

$$p(\underline{L}) = \frac{2^c}{F(n, c, \underline{I})} L_{cc}^{2n-1} \dots L_{11}^{2n-2c+1} \exp \left[ -\text{Tr} (\underline{L}^t \underline{L}) \right] \quad (A-13)$$

and after,

$$p(L_{11}, \hat{\underline{\Gamma}}, \underline{M}) = \frac{2^c}{F(n, c, \underline{I})} L_{11}^{2n-2c+1} e^{-L_{11}^2} |\underline{M}|^2 L_{cc}^{2n-1} \dots L_{22}^{2n-2c+3} e^{-\text{Tr} \left[ (\underline{I} + \hat{\underline{\Gamma}} \hat{\underline{\Gamma}}^t) M^t M \right]} \quad (A-14)$$

Equation (A-6) of Lemma 2 is used to integrate over  $\underline{M}$ . The resulting marginal density of  $(L_{11}, \hat{\underline{\Gamma}})$  is

$$p(L_{11}, \hat{\underline{\Gamma}}) = \frac{2 L_{11}^{2n-2c+1} e^{-L_{11}^2}}{F(n, c, \underline{I})} F \left( n+1, c-1, \left[ \underline{I} + \hat{\underline{\Gamma}} \hat{\underline{\Gamma}}^t \right]^{-1} \right) \quad (A-15)$$

The final transformation  $\beta = \frac{L_{11}^2}{n}$  obtains the density for  $\beta$  and  $\hat{\underline{\Gamma}}$ :

$$p(\beta, \hat{\underline{\Gamma}}) = \left[ \frac{n^{n-c+1} \beta^{n-c} e^{-n\beta}}{(n-c)!} \right] \left[ \frac{n!}{\pi^{c-1} (n-c+1)!} \frac{1}{(1 + |\hat{\underline{\Gamma}}|^2)^{n+1}} \right] \quad (A-16)$$

$$= p(\beta) p(\hat{\underline{\Gamma}}) \quad \text{for } \beta \geq 0, \hat{\underline{\Gamma}} \text{ complex}$$



---

Thus,  $\beta$  and  $\hat{\underline{\Gamma}}$  are statistically independent.

The density for the quantity

$$\alpha = 1 + \hat{\underline{\Gamma}}^t \hat{\underline{\Gamma}} = 1 + |\hat{\underline{\Gamma}}|^2$$

can now be determined as

$$p(\alpha) = \frac{n!}{(n-c+1)! (c-2)!} \frac{(\alpha - 1)^{c-2}}{\alpha^{n+1}} \quad \text{for } \alpha \geq 1, n \geq c \quad (A-17)$$



---

**APPENDIX B**  
**SUMMARY OF NOTATION**



---

## APPENDIX B

### SUMMARY OF NOTATION

$\underline{x}$  c-dimensional complex Gaussian vector representing Fourier transforms of c channels in an array

$\underline{x}^{(1)}$  First component of  $\underline{x}$ ; principal channel being estimated by the other c-1 channels

$\underline{\Sigma}$  c by c positive definite covariance matrix of channels;  
 $\underline{\Sigma} = E\{\underline{x} \underline{x}^t\}$

$\hat{\underline{\Sigma}}$  Sample covariance of n independent samples  $\underline{x}_k$  ( $k = 1, \dots, n$ ) from the array; estimate of true covariance  $\underline{\Sigma}$

$\underline{f}$  (c-1) dimensional filter used to weight the c-1 channels to form an estimate of  $\underline{x}^{(1)}$

$\underline{f}_o$  Optimum filter; minimum mean-square-error filter;  
Wiener filter

$\hat{\underline{f}}$  Estimated filter; filter designed on basis of n independent observations (samples) of  $\underline{x}$  with unknown  $\underline{\Sigma}$

$\sigma_o^2$  Performance of Wiener filter; minimum attainable mean-square error

$\sigma^2$  True performance of estimated filter

$\hat{\sigma}_o^2$  Least-squares error of estimated filter operating on n design points; regression error; estimated performance of estimated filter

$\alpha$  Ratio of performance  $\frac{\text{estimated-filter error}}{\text{optimal-filter error}} = \frac{\sigma^2}{\sigma_o^2}$

$\beta$  Ratio of  $\frac{\text{regression error of estimated filter}}{\text{optimal-filter error}} = \frac{\hat{\sigma}_o^2}{\sigma_o^2}$

**UNCLASSIFIED**

Security Classification

**DOCUMENT CONTROL DATA - R&D**

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Texas Instruments Incorporated Science Services Division P. O. Box 5621, Dallas, Texas 75222		2. REPORT SECURITY CLASSIFICATION Unclassified
		3. GROUP
4. REPORT TITLE <b>STATISTICS GOVERNING THE DESIGN AND PERFORMANCE OF NOISE-PREDICTION FILTERS - ADVANCED ARRAY RESEARCH - SPECIAL REPORT NO. 3</b>		
5. DESCRIPTIVE NOTES (Type of report and inclusive dates) <b>Special</b>		
6. AUTHOR(S) (Last name, first name, initial) Burg, John P.      Cruise, T. J.      Booker, Aaron H.		
7. REPORT DATE 8 September 1967	7a. TOTAL NO. OF PAGES 46	7b. NO. OF REFS 2
8a. CONTRACT OR GRANT NO. F33657-67-C-0708-P001	8b. ORIGINATOR'S REPORT NUMBER(S) _____	
8c. PROJECT NO. VELA T/7701	8d. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) _____	
10. AVAILABILITY/LIMITATION NOTICES This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of Chief, AFTAC.		
11. SUPPLEMENTARY NOTES ARPA Order No. 624 ARPA Program Code No. 7F10	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency Department of Defense The Pentagon, Washington, D. C., 20301	

13. ABSTRACT

In designing a digital multichannel filter from a limited sample of noise, a highly important parameter,  $\alpha$ , is defined as the true mean-square error of the estimated filter (i. e., the average long-term performance of the filter obtained from the noise sample) divided by the true mean-square error of the optimum filter.

The value of  $\alpha$ , which is equal to or greater than one, is not known before or after an experiment since the true covariance of the data is required to calculate its value; however, the probability density of  $\alpha$  turns out to be invariant with respect to the true covariance and depends only on the amount of data and the number of channels in the filter. Thus, one can determine before collecting any data how long a sample is needed in order to design a filter which is within 1 db (for example) of optimum with 90-percent confidence. A second similar parameter,  $\beta$ , defined as the estimated mean-square error of the estimated filter (i. e., the regression error) divided by the true mean-square error of the optimum filter is highly useful in deciding the reliability of the apparent effectiveness of the designed filter.

The probability densities of  $\alpha$  and  $\beta$  are derived for the Gaussian assumption and graphs useful in experiment design are presented in this report.

## UNCLASSIFIED

## Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Advanced Array Research						
Design of Noise Prediction Filters						
Performance of Noise Prediction Filters						
Mean-Square Estimation Formulation						
Invariance Properties of $\alpha$ and $\beta$						
Probability Densities of $\alpha$ and $\beta$						
Confidence Levels						

## INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those

imposed by security classification, using standard statements such as:

- (1) "Qualified requesters may obtain copies of this report from DDC."
- (2) "Foreign announcement and dissemination of this report by DDC is not authorized."
- (3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through \_\_\_\_\_."
- (4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through \_\_\_\_\_."
- (5) "All distribution of this report is controlled. Qualified DDC users shall request through \_\_\_\_\_."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.